AI-Powered Smart Glasses for the Blind and Visually Impaired

M.Ananthi[1,], R.Bharathi[2], M.Gayathri[3], G.Gokul[4], M.Sivakumar[5], B.Vaishnavi[6,]

[1,2]*Assistant Professor, Department of Computer Science andEngineering,Cheran College Of Engineering, Karur,Tamilnadu,India.*

[3,4,5,6] *U.G.Student,Department of Computer Science and Engineering,Cheran College Of Engineering, Karur,Tamilnadu,India.*

[1]*rithu.ananthi@gmail.com,*[2]*bharathimkce@gmailcom,*
[3]*gayathrimurugan810@gmail.com,*[4]*gokul302001@gmail.com,*[5]*viratsiva342@gmail.com,*[6]*vaishnavibaskaran060@ gmail.com*

Abstract

Vision is one of the most crucial human senses, because it helps us perceive our surroundings. However, millions of individuals throughout the globe suffer from vision loss. They are having difficulty navigating their everyday lives because they are unable to detect impediments in their environment, and one of their biggest challenges is identifying people. Other than automation, object detection is used in a variety of applications that have yet to be fully explored. This project includes one such application that employs detection to assist visually impaired individuals in identifying items ahead of them for safe navigation, as well as a face recognition system with aural output that may help visually impaired people recognise known and unfamiliar people. Speakers would provide them with voice-based assistance. We used a deep learning-based Faster Region-Convolutional Neural Network (Faster R-CNN) to identify and recognise humans and objects in the environment in this study. The Faster Region Convolution Neural Network technique processes and classifies the picture taken by the camera. The audio jockey receives the detected picture as an audio input. As a result, this model aids visually impaired persons in a more comfortable manner than white canes.

Keywords: VisualImpairment,Objectdetection,Faster Region Convolutional Neural Network,Audio Jockey.

1. INTRODUCTION

"Visual impairment" is a wide phrase that refers to any degree of visual loss that interferes with a person's ability to carry out everyday activities.

Types of Vision Impairment

Vision impairments are defined differently in different nations. Visual impairment is classified by the World Health Organization (WHO) based on two factors: visual acuity, or clarity of vision, and visual fields, which is the region from which you can detect visual information when your eyes are fixed and gazing directly at an object.

The SnellenChart is a visual acuity test. Two values are used to compute your visual acuity. The first value represents the distance between the reader and the chart.
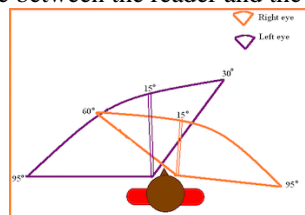

Figure.1.Snellen Chart using for calculating distance to opposite object

The second figure is how far a person with normal vision would have to stand from an item to view what you saw at 20 feet. A visual acuity of 20/80, for example, indicates you can see the chart from 20 feet away as well as someone who can read it from 80 feet away. In other words, you won't be able to see what a person with normal vision would see from 80 feet away until you get closer to just 20 feet away. This illustration depicts how the visual fields are formed by crossing the eyes [1]-[5].

2. RELATED WORK

Apart from providing a helping hand, assistive technology adapted with AI, material engineering, and robotics offers a number of ground-breaking instances that aid the physically crippled, intellectually challenged, and disabled.

• Visual Aids Powered by AI

Microsoft's Seeing AI software for persons with vision impairment has already been published. This software allows the user to hold their phone up to a person and have the phone describe their appearance, hair colour, age, if they seem happy or unhappy, and so on. You can learn what a product is, when it expires, and other information by pointing your phone at it. In addition, the software can read texts and detect structure components like paragraphs, headers, and lists.

• Glasses with Intelligence:

Smart glasses are currently in the early stages of development and are projected to enter the market in 2020. Smart glasses are equipped with augmented and virtual reality contrast lenses, allowing individuals to use them instead of glasses or contact lenses. These glasses assist people concentrate on a specific point of focus by allowing them to converge and diverge light beams. Both neuroscientists and computer vision experts at the University of Oxford are working on this technology to make it the greatest of its type.[6]-[10]

• Hearing Aids with Cognitive Functions

Hearing aids have been around for a long time. However, cognitive hearing aids are a dream accessory for everyone. These sophisticated intelligent hearing aids may tune in with brain waves to understand what a someone wants to hear at a certain moment depending on their mood swings, in addition to assisting with hearing. Its capacity to read people's minds and respond accordingly is what sets it apart. This implies that whether numerous individuals are speaking at the same time or if another kind of voice is dominant, the gadget may monitor and adjust the user's brain and silence all other sounds except what the user wishes to hear.

Text To Sign

Sign language is mostly used by those who are deaf or hard of hearing. However, not everyone is fluent in this language, since many people throughout the world are unaware of half of the signals used to communicate. As a result, firms all around the globe are developing prototypes that can transform sign language into text or speech that the communicator and receiver can comprehend. The gadget has a 3D camera that tracks, analyses, and interprets a person's signing motion in order to transmit it to others. This device's beta version is now on the market, and it can accurately understand messages with 98 percent accuracy. At both a personal and professional level, the combination of artificial intelligence, machine learning, and robots is changing assistive technology. The improvements in this technology are allowing the physically and intellectually challenged to live a normal life, as well as offering healthcare facilities with sophisticated systems to maximise patient treatment as soon as possible. Collaboration between human and machine intelligence may improve healthcare results in the future.

Machine learning (ML) is a powerful method to artificial intelligence (AI) that is aimed to create patterns from data and is becoming more prevalent in our everyday lives. ML procedures are designed to anticipate how various entities will behave in the future, enabling choices to be made based on those predictions.

3. PROPOSED SYSTEM

The project's planned solution aims to build and construct a Smart Electronic Glass that would aid visually challenged persons in identifying faces and things.

Face Recognition and Object Detection is a computer technique that deals with finding instances of semantic items of a certain class (such as individuals, buildings, or vehicles) in digital photos and videos. Face recognition and object detection are two disciplines that have seen a lot of progress. Face identification (used by Facebook to distinguish individuals), tumour detection (used in medical disciplines), and other applications utilise it. Object recognition problems have gotten substantially easy and efficient with the introduction of deep learning in computer vision.

Deep learning models outperform prior computer vision algorithms in terms of accuracy, time consumption, complexity, and overall performance. Deep learning outperformed previous computer vision approaches for object identification, leading to widespread use of deep learning models.

One of the most effective object identification algorithms (deep learning) is:

1. RCNN (Region-based Convolution Neural Network)
2. Rapid RCNN
3. Improved RCNN

R-CNN uses selective search to extract a number of areas from a given picture, and then evaluates whether any of these boxes contain an item. We first extract these areas, then utilise CNN to extract particular characteristics for each region. These characteristics are then utilised to detect things. Unfortunately, because of the various phases involved in the process, R-CNN becomes rather sluggish.
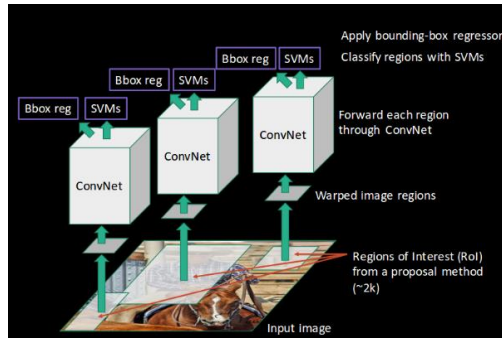
Figure.2.Fast Region-convolution Neural Network (Fast-RCNN)

In contrast, Fast R-CNN sends the full picture to ConvNet, which creates areas of interest (instead of passing the extracted regions from the image). It also employs a single model that collects features from the areas, classifies them into distinct classes, and outputs the bounding boxes, rather than three independent models (as shown in R-CNN) is shown in fig.2

All of these stages are completed at the same time, making it quicker than R-CNN. When applied to a big dataset, however, R-CNN is not quick enough since it employs selective search to extract the areas. Is shown in fig.3.
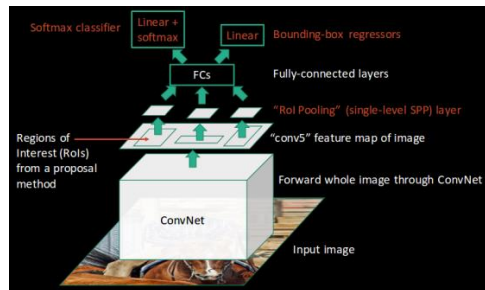

Figure.3.Faster Region-Convolution Neural Network(Faster-RCNN)

4. Results and Discussion

The sensor captures the picture of the opposing item, similar to a human face detection module. The image of a face is classed as a known or unknown person or an object. The training classified result and the test Live Cemara Captured Classified file are matched in this module. If a process data is triggered during processing, speech synthesis is utilised to warn the user, for example, by creating the word "stop." The audio for "Hi Ram" has been released is shown in 4 and 5.
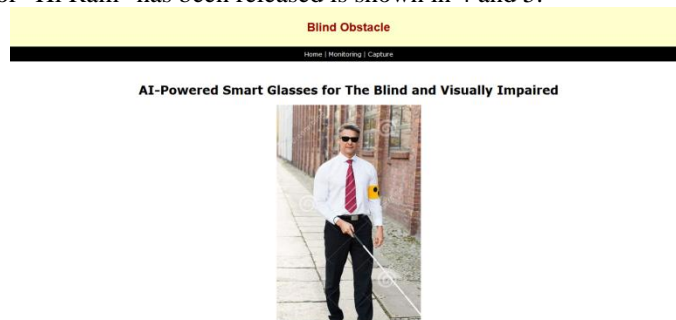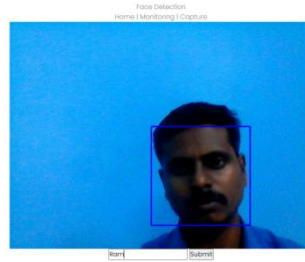

Figure.4.SmartClassforblindpeople

Figure.5.Image intentifing for known or unknown person

5. Conclusion

The item on display is a smart glass that combines the capabilities of a machine vision and obstacle detection and identification sensor. It can be easily promoted and made accessible to those who are visually impaired. It would also aid in the prevention of future injuries. Smart devices may be easily carried, and the system camera can monitor things and faces in the area and show them in audio format. Each model denotes a distinct job or mode. The user may execute the chosen task separately from the other tasks. The system's design, functioning mechanism, and principles, as well as certain experiment outcomes, were addressed. Allow visually impaired persons to engage more intimately with others around them without worry of their vision becoming clouded and hazy.

6. Future Work

Although it is still a prototype, our system represents a promising avenue for future research aimed at enhancing the spatial awareness of visually impaired people traveling in unfamiliar environments.

7. References

1. F. Catherine, ShiriAzenkot, Maya Cakmak, "Designing a Robot Guide for Blind People in Indoor Environments," ACM/IEEE International Conference on Human-Robot Interaction Extended Abstracts, 2015.
2. H. E. Chen, Y. Y. Lin, C. H. Chen, I. F. Wang, "Blindnavi: a mobile navigation app specially designed for the visually impaired," ACM Conference Extended Abstracts on Human Factors in Computing Systems, 2015.
3. K. W. Chen, C. H. Wang, X. Wei, Q. Liang, C. S. Chen, M. H. Yang, and Y. P. Hung, "Vision-based positioning for Internet-of-Vehicles," IEEE Transactions on Intelligent Transportation Systems, vol. 18, no.2, pp. 364–376, 2016.
4. M. Cordts, M. Omran, S. Ramos, T. Rehfeld, M. Enzweiler, R. Benenson, U. Franke, S. Roth, and B. Schiele, "The Cityscapes Dataset for Semantic Urban Scene Understanding," IEEE Conference on Computer Vision and Pattern Recognition, 2016.
5. J. Ducasse, M. Macé, M. Serrano, and C. Jouffrais, "Tangible Reels: Construction and Exploration of Tangible Maps by Visually Impaired Users," ACM CHI Conference on Human Factors in Computing Systems, 2016.
6. J. Engel, T. Schops, and D. Cremers, "LSD-SLAM: Large-scale direct monocular SLAM," European Conference on Computer Vision, 2014.
7. S. Gilson, S. Gohil, F. Khan, V. Nagaonkar, "A Wireless Navigation System for the Visually Impaired," Capstone Spring, 2015.
8. J. Guerreiro, D. Ahmetovic, K. M. Kitani, and C. Asakawa, "Virtual Navigation for Blind People: Building Sequential Representations of the Real-World," International ACM SIGACCESS Conference on Computers and Accessibility, 2017.
9. Kendall, M. Grimes, and R. Cipolla, "PoseNet: a convolutional network for real-time 6-DOF camera relocalization," International Conference on Computer Vision, 2015.
10. Kendall, and R. Cipolla, "Geometric loss function for camera pose regression with deep learning," International Conference on Computer Vision, 2017.
11. R.Bharathi, T.Abirami," Energy efficient compressive sensing with predictive model for IoT based medical data transmission", Journal of Ambient Intelligence and Humanized Computing, November 2020, https://doi.org/10.1007/s12652-020-02670-z
12. R.Bharathi, T.Abirami," Energy Efficient Clustering with Disease Diagnosis Model for IoT based Sustainable Healthcare Systems", Sustainable Computing: Informatics and Systems, 23 September 2020, https://doi.org/10.1016/j.suscom.2020.100453